DOCUMENT RESUME

ED 271 496 TM 860 405

AUTHOR Thathong, Ngamnit; Kruawan, Preecha
TITLE Self-Scoring Flexilevel Test.
PUB DATE [Jan 85]
NOTE 12p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adaptive Testing; Feasibility Studies; Foreign Countries; High Schools; Mathematics Achievement; Mathematics Tests; *Scoring; Secondary School Mathematics; *Self Evaluation (Individuals); *Student Reaction; Test Construction; Test Theory; *Test Validity
IDENTIFIERS *Self Scoring Tests; Thailand

ABSTRACT

The feasibility of a self-scoring flexilevel test was investigated in terms of practical effectiveness and criterion related validity. The test was administered to over 2,000 students studying secondary school mathematics in Thailand. The study was conducted, and the test administered, in two phases: test development and evaluation. The test was designed to measure a single ability and the items were multiple choice. A large item bank was used in order to spread item difficulty values over a wide range and to have a constant distance between item difficulty values. Results indicated that the self-scoring flexilevel test had a criterion related validity of .7238. This was statistically significant and different from the validity coefficient (.6109) of the comparative conventional test. Results of an attitude measure indicated that the students had a favorable impression of the self-scoring flexilevel test. It was concluded that this type of test is effective and practical. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED271496

Self - scoring flexilevel test

by

Ngamnit Thathong and Preecha Kruawan

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

N. Thathong

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 860 405

Self-scoring flexilevel test

Ngamnit Thathong and Preecha Kruawan

Khon Kaen University, Thailand

Abstract

This study investigated the feasibility of the self-scoring flexilevel test in terms of practical effectiveness and criterion-related validity in topics of ratio and percentage in mathematics course of Mathayom Suksa II. This study was divided into two phases. Construction of the self-scoring flexilevel test was in phase I, while phase II was to evaluate the self-scoring flexilevel.

The criterion-related validity coefficient of the self-scoring flexilevel test was .7238 which was statistical significant difference from the validity coefficient of the conventional test (.6109) at the .01 level. Furthermore, the students had good impression and good attitude in dealing with the self-scoring flexilevel test.

The results show that the self-scoring flexilevel test is effective according to the validating criteria and practicality.

Self-scoring flexilevel test

Ngamnit Thathong and Preecha Kruawan

Khon Kaen University

The purpose of this study was to investigate the feasibility of using the self-scoring flexilevel test in school testing. This was accomplished by constructing the conventional test in topics of Ratio and Percentage in mathematics course of Matayom Suksa II, modifying the directions, the test booklets, and the answer sheets to become a flexilevel test (Lord, 1971 a, b), and then administering them to the examinees

According to Ausubel (1968), motivation is greatest when the tasks are appropriate to student's ability and never too easy or too difficult. Thus, individualizing tests might be expected to enhance a student's motivation to complete it.

It is well known that the most effective test items are those for which the item difficulty values are appropriate for the average ability level consisting of students who know half of the items. The conventional test previously referred to, a multiple choice test, requires that all students answer the same set of test items. This has caused less effective measurement for the examinee whose ability deviates from the average. That is, items which appear to be overly easy for some examinees appear to overly difficult for others.

Adaptive ability measures test (tailored, branced, individualized, flexilevel, for example) have attempted to provide test items which are best suited for the examinee's ability level by using the preceeding item response to determine the next item to which the examinee is to respond. In adaptive testing, a large pool of items must be available to start with, and the statistical characteristices of these items must be known from earlier testing. For this reason adaptive testing has not often been used in actual testing situations.

Method

Subjects

The subjects were selected randomly by cluster sampling method from the students in the Secondary School Group I of Kalasin province, Department of General Education, Ministry of Education. In phase I, the first tryout sample consisted of 1,160 students in the second semester of Mathayom Suksa II, 1981. 476 students in the first semester of Mathayom Suksa III,1982, was the second tryout sample. In phase II, the sample was 535 students in Mathayom Suksa II of the second semester, 1982.

Instruments

In order to construct a flexilevel test four conditions should be met. First, the test should measure a single ability. Second, the item difficulty values should be spread evenly over the widest possible range. Third, the difference between adjacent item difficulty values (step-size) should be constant. In order to meet the second and third condition, a large item pool is required. Last, the items should be the multiple-choice type.

In phase I, 58 items out of 120 items in the first four tryout subtests were selected according to the discrimination index of .295 for construction the second tryout test. Fifty-three items from the second tryout test were selected according to the discrimination index of .295 and formed to be a self-scoring flexilevel and a conventional tests. Finally, the self-scoring flexilevel test consisted of 35 items. The item difficulties ranged from .18 to .87 with the average was .50. The mean of differ in difficulty at successive stage was .02. Item discriminations ranged from .34 to .73 with the average was .51.

In phase II, the instruments were the self-scoring flexilevel test, the conventional test, and the students' opinionnaire toward the feasibility of the self-scoring flexilevel test.

In producting the flexilevel test booklets, answer sheets were prepared by modifying Lord's answer sheets as shown in Figure 1.

Insert Figure 1 about here

The direction formed the first page of the test booklet with the starting (median difficulty) item given on the top of the second page. Opening the test booklet from the second to the fifth pages exposed the seventeen easy items in the left colum and the seventeen hard items in the right column of the same page.

Each examinee was asked to use tip pen open the four boxes corresponding to his selection of the answer to an items. This caused the letters ก or ข to apear. When an item was answered and the word ก appeared, the examinee went to answer the next unanswered ก items. when the word ข appeared, the examinee went to answer the unanswered ข items. After responding to each item the examinee was directed to check the box adjacent to the item response on his answer sheet (the center column of boxes in Figure 1). When a check showed in each of the black boxes, the examinee stopped as he had answered $\frac{n+1}{2}$ or 18 ก and ข items.

The scoring in a flexilevel test can be done in the following steps.

1. If an examinee's last answer was correct (a word ข appeared), his number-right score equals the next serial unanswered ข items.

2. If the last item response was incorrect (the word ก appeared), the item number of the next serial unanswered ก item is the examinee's number-wrong score. This value must be subtracted from $\frac{(n+1)}{2}$ to get his number-right score. According to Lord (1971 a), this examinee should also receive a bonus one-half point as he has outperformed all examiness taking the same set of items that ended with a corfect answer.

Statistical Analysis

Data were analized to determine the relationships between the student ability levels and test scores derived from the self-scoring flexilevel and the conventional tests. In addition, basic statistics and percentage were computed to determine the efficiency of the self-scoring flexilevel test.

Results

Table 1 indicates the basic statistics of test scores form both test. It can be seen from Table 1 that scores derived from the flexilevel test has the smaller variability which can be interpreted to mean that the students' ability are closed together. Table 1 also shows that a mean of the flexilevel test scores is greater than a mean of the conventional test scores.

Insert Table 1 about here

The correlation coefficients between the self-scoring flexileval test the conventional test and the students' ability (θ) are indicated in

Table 2 which each coefficient was significant difference from zero at the .01 level. It is apparent that both the self-scoring flexilevel and the conventional tests are valid tests but there is significant difference between the validity coefficients for the self-scoring flexilevel test and the conventional test as indicates in table 3.

Insert Table 2 about here

Insert Table 3 about here

Table 4 shows the basic statistics of the students' opinion toward the self-scoring flexilevel test based on 535 subjects in five categories.

Insert Table 4 about here

As we can see from Table 4 that the students show positive attitude toward the self-scoring flexilevel test.

Summary and conclusions

The following findings of this study should be of interest to those who are involved with flexilevel test in school testing.

First, the self-scoring is more effective than the conventional test according to the criterion-related validity.

Second, the self-scoring flexilevel is feasible in using with school testing.

References

Ausubel, D. P. Educational psychology : A cognitive review. New York : Holt, Rinehart, and Winton, 1968.

Dick and Hagerty. Topics in measurement : Reliability and validity. New York : McGraw-Hill, 1971.

Lord, F. M. Some test theory for tailored testing. In W. H. Holtzman (Ed), Computer-assisted instruction, testing, and guidance. Chapter VIII. New York : Harper & Row, 1970.

Lord, F. M. The self-scoring flexilevel test. Journal of Educational Measurement, 8, 147-151, 1971 (a)

Lord, F. M. A theoretical study of the measurement effectiveness of flexilevel tests. Educational and Psychological Measurement, 31, 805-813, 1971 (b).

Nunnally, J. C. Psychometric theory. New York : McGraw-Hill, 1969.

Olivier, Philippe. An evaluation of the self-scoring flexilevel tailored testing model. Doctoral dissertation. The Florida State University, 1974.

BEST COPY AVAILABLE

Median difficult item									
(0)	1	2	3	4					
ก-1	(easier than starting item)					ข-17	(hardest item on test)		
ก -2	1	2	3	4		ข -16	1	2	3	4
ก -3	1	2	3	4		ข -15	1	2	3	4
ก -4	1	2	3	4		ข -14	1	2	3	4
ก -5	1	2	3	4		ข -13	1	2	3	4
ก -6	1	2	3	4		ข -12	1	2	3	4
ก -7	1	2	3	4		ข -11	1	2	3	4
ก -8	1	2	3	4		ข -10	1	2	3	4
ก -9	1	2	3	4		ข -9	1	2	3	4
ก -10	1	2	3	4		ข -8	1	2	3	4
ก -11	1	2	3	4		ข -7	1	2	3	4
ก -12	1	2	3	4		ข -6	1	2	3	4
ก -13	1	2	3	4		ข -5	1	2	3	4
ก -14	1	2	3	4		ข -4	1	2	3	4
ก -15	1	2	3	4		ข -3	1	2	3	4
ก -16	1	2	3	4		ข -2	1	2	3	4
ก -17	(easiest item on test)					ข -1	(harder than starting item)			
ก						ข				

Figure 1 Layout the flexilevel answer sheet used in this study.

Table 1

Basic statistics of scores derived form the self-scoring flexilevel and the conventional test.

	Self-scoring flexilevel test	Conventional test
Mean	9.884	8.321
S.D.	3.682	4.015
Maximum	18.00	18.00
Minimum	2.00	1.00
Range	10.00	17.00

Table 2

Pearson Product Moment correlation coefficients between scores from the self-scoring flexilevel, the conventional test, and the students' ability level.

	Self-scoring flexilevel test	Conventional test
Students' ability level	.7238**	.6109**
Conventional test	.7529**	

** $p < .01$

Table 3

Criterion-related coefficients between the self-scoring flexilevel test and the conventional test.

Self-scoring flexilevel test	Conventional test	t
.7238	.6109	5.0023**

** $p < .01$

Table 4

Mean and standard deviation of opinion scores toward the self-scoring flexilevel test.

Category	Mean	Standard deviation
1. Direction and administration	3.4676	1.1184
2. Test booklets and answer sheets	3.8151	1.1119
3. Perception of item difficulties	3.9774	1.0459
4. Motivation and anxiety	3.7751	1.1708
5. General opionion	3.2909	1.2900

BEST COPY AVAILABLE